Title of the Project : MALWARE DETECTION AND FAMILY CLASSIFICATION WITH DYNAMIC FEATURES USING ML

Name of the Students : HARIKARAN K M , IMMANUVEL V

Register Number(s) : 211423104199 ,211423104228

Name of the Guide : ALWIN INFANT P

**ABSTRACT**

The increasing complexity and sophistication of malware pose a significant threat to cybersecurity. Traditional detection approaches, such as static signature-based methods and heuristic disassembly, often fail against modern malware because attackers exploit techniques like obfuscation, encryption, and polymorphism to evade detection. This creates a pressing need for more adaptive and intelligent detection frameworks.

To address these challenges, this project introduces MalwareClassifier, a two-stage ensemble machine learning framework. In the first stage, a stacked ensemble of Logistic Regression, Random Forest, and XGBoost is employed to perform binary classification, distinguishing benign samples from malicious ones. This layered approach ensures greater robustness and accuracy by leveraging the complementary strengths of different classifiers.

In the second stage, malware samples identified in Stage 1 undergo fine-grained analysis through a family classification model. Here, ensemble methods, including Logistic Regression, Random Forest, and XGBoost, are combined to categorize malware into families such as Trojans, spyware, and ransomware. This hierarchical approach not only enhances detection performance but also generates actionable threat intelligence, which is crucial in cybersecurity operations.

Extensive preprocessing techniques such as label encoding, feature scaling, and missing value imputation are integrated into the pipeline, making it highly suitable for structured datasets like CIC-MalMem-2022. Experimental results demonstrate that the proposed system outperforms traditional static-analysis-based methods, achieving higher accuracy, better generalization, and resilience against evasion strategies. This highlights its potential for real-world malware detection and defense systems.